View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

# Adding 'splice' to protein engineering

Mande Holford and Tom W Muir*

Address: Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

*Corresponding author.
E-mail: muirt@rockvax.rockefeller.edu

**Structure** 15 August 1998, **6**:951–956
http://biomednet.com/elecref/0969212600600951

© Current Biology Publications ISSN 0969-2126

The field of protein engineering has always attracted a multidisciplinary cast of characters ranging from biologists to organic chemists. Consequently, a variety of different protein engineering techniques have been described which encompass the purely synthetic and the purely biosynthetic. Of these, heterologous expression of recombinantly cloned genes is by far the most commonly employed route to engineered proteins — in part because the approach is readily accessible to most researchers. Although site-directed mutagenesis is facilitated by the 'replicative fidelity and efficiency of living organisms' [1], this of course also constrains the approach to the naturally occurring amino acids. This limitation is largely overcome by elegant chemistry-driven approaches, such as unnatural amino acid mutagenesis [2] and protein total or semi-synthesis via the chemical [3–9] or enzymatic [10–13] ligation of unprotected peptides. These approaches are, however, technically demanding, a feature which has prevented their widespread use by the biological community. Recently, a novel protein semi-synthesis approach was reported [14–16] that harnesses the phenomenon of protein splicing (for review see [17]), and in the process renders proteins harboring unnatural building blocks accessible to most researchers. This short review describes the genesis of this methodology, termed 'expressed protein ligation' [14,15], its applications thus far and its future potential in protein engineering.

### The development of expressed protein ligation

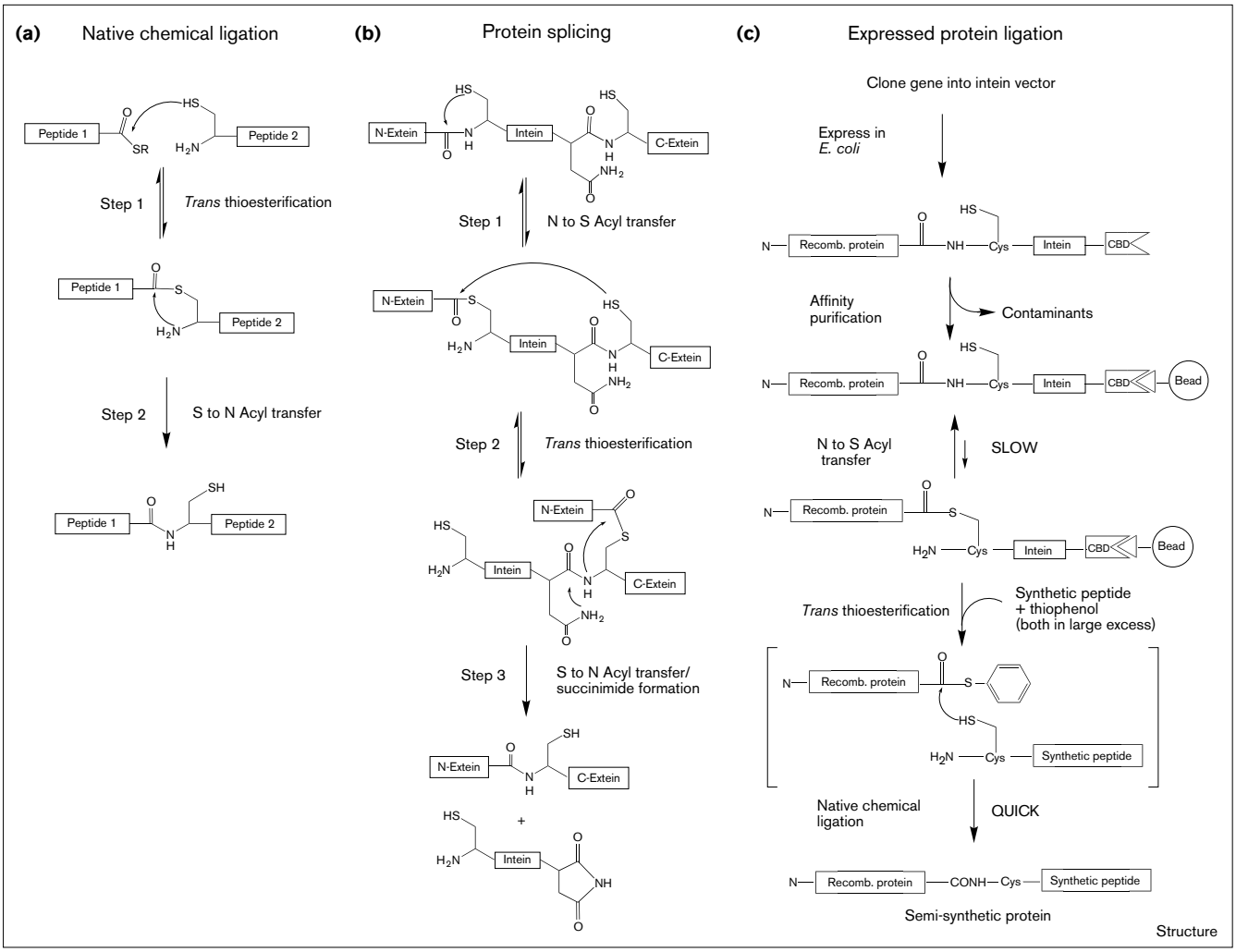
Expressed protein ligation (EPL) allows synthetic peptides to be chemically ligated to the C terminus of recombinant proteins through a normal peptide bond. The approach arises from the unexpected convergence of two areas of research which, at first glance, appear unrelated: protein total synthesis [18] and protein self-splicing [17]. The total chemical synthesis of proteins has in recent years been dominated by the use of powerful fragment condensation approaches that allow fully unprotected peptide building blocks to be regioselectively pieced together in aqueous solution [19]. Unlike full-length proteins, peptide fragments can be efficiently assembled by stepwise solid-phase peptide synthesis (synthetic polypeptides of up to ~50 residues are now readily accessible [20]). Moreover, by using unprotected peptides these modern ligation approaches overcome the solubility and characterization problems associated with classical convergent synthesis (reviewed in [21]).

Among the various chemical and enzymatic ligation methods in use, so-called 'native chemical ligation' [22] has proven the most general route to fully synthetic proteins [22–27]. As illustrated in Figure 1a, native chemical ligation relies on the highly efficient reaction that occurs between a peptide possessing an N-terminal cysteine residue (peptide 2; Figure 1a) and a second peptide possessing an $^{\alpha}$thioester moiety (peptide 1; Figure 1a). Following an initial intermolecular, and important chemoselective reaction, the resulting thioester-linked intermediate spontaneously rearranges via an S→N acyl transfer to give the final amide-linked product. As was pointed out in the initial description of the technique [22], the C-terminal segment could be expressed using recombinant DNA methods; provided the polypeptide contained an N-terminal cysteine, it could be used in native chemical ligation. Indeed, this semi-synthetic version of native chemical ligation has been successfully used to attach an ethylenediamine-tetraacetic acid (EDTA) analog to the N terminus of a chimeric heterodimer of the transcription factors c-Jun and c-Fos [28].

Although biotechnology provides the necessary tools required to generate N-terminal cysteine-containing proteins, the same has not been true for the production of recombinant protein $^{\alpha}$thioesters. Thus, native chemical ligation could not be fully interfaced with recombinant protein biotechnology. It is at this point that we must introduce protein splicing into our story, as this phenomenon provides the final piece of the puzzle. Protein splicing, first detected in yeast nearly a decade ago [29], is a post-translational event in which a proprotein undergoes an intramolecular rearrangement resulting in the extrusion of an internal sequence (termed an intein [30]) and the concomitant joining together of the flanking sequences (exteins) to form an edited version of the original protein. Unlike RNA splicing, which requires the help of an elaborate ribonuclear protein complex, protein splicing is an intrinsic process and can take place *in vitro* using purified proproteins [31]. This has allowed the mechanism of the process to be studied in some detail using both mutagenesis and protein chemistry based techniques [31–33].

Our current understanding of the mechanism of protein splicing is summarized in Figure 1b. The first step involves an N→S or N→O acyl shift in which the N-extein is

**Figure 1**




Native chemical ligation and protein splicing provide the basis for a new protein semi-synthesis method. The three different techniques discussed in the text are illustrated: **(a)** native chemical ligation, **(b)** protein splicing and **(c)** expressed protein ligation.

transferred to the sidechain sulfhydryl/hydroxyl group of the first residue in the intein — a cysteine, serine or threonine residue is always located at this position. This rearrangement would appear to be energetically unfavorable and is essentially the reverse of the final step in native chemical ligation (Figure 1a). Although the driving force(s) behind this event is (are) still the subject of study, it is interesting to note that in the recently solved crystal structure of the GyrA intein, the amide bond formed between the final residue in the N-extein and the first residue in the intein is observed to be in the higher energy *cis* conformation [34]. It is possible that this unusual stereochemical constraint helps push the equilibrium of the system towards the ester/thioester side. Given that all inteins are likely to share a common fold, it is intriguing to speculate that this type of stereochemical 'leg-up' may be a general feature of protein splicing.

After the initial N→(S/O) rearrangement there follows a *trans* esterification step in which the N-extein is transferred to the sidechain of a second conserved cysteine, serine or threonine residue, this time located at the intein–C-extein junction. In the next step, the amide bond at the intein–C-extein junction is broken as a result of succinimide formation involving the conserved C-terminal asparagine residue within the intein. In the final step of the process, a peptide bond is formed between the N-extein and C-extein following an (S/O)→N acyl shift — identical to the final step in native chemical ligation.

The uncanny chemical similarities between native chemical ligation (Figure 1a) and protein splicing (Figure 1b) raises an interesting question: is it possible to use protein splicing as a route to the formation of recombinant protein $^{\alpha}$thioesters for use in native chemical ligation? This

question is further highlighted following the recent clever demonstration by Xu and coworkers that protein splicing can be halted in midstream (presumably at the thioester /ester intermediate stage) by mutation of the key asparagine residue in the intein to an alanine residue [35]. Indeed, an engineered intein of this type now forms the basis of a commercially available protein expression/purification strategy (New England Biolabs).

Now with all the pieces in place, recent months have seen the publication of a series of papers [14–16] that pull together protein splicing and native chemical ligation into a general new protein semi-synthesis approach. The principle of the method, referred to as expressed protein ligation (EPL) [14,15], is shown in Figure 1c. In the first step, the protein or protein fragment of interest is expressed as an intein–CBD (chitin binding domain) fusion protein, where the CBD allows the protein of interest to be affinity purified using chitin beads. Because an engineered intein is being used, protein splicing is unable to proceed to completion and is thus stalled at the stage involving the thioester-linked intermediate. Consequently, exposure of the immobilized fusion protein to an aqueous solution containing of an N-terminal cysteine-peptide and a thiol reagent (e.g. thiophenol) results in native chemical ligation of the peptide to the protein. The nature of the thiol agent used is of key importance to this process as it dictates the reactivity of the protein $^{\alpha}$thioester generated *in situ* during the reaction. Thus, thiophenol is a good choice as phenyl $^{\alpha}$thioesters are known to be extremely reactive in native chemical ligations [25].

Although the three reports to date all use more or less the same procedure shown in Figure 1c, they each illustrate a different potential application of EPL. The first study makes use of the commercially available *Saccharomyces cerevisiae (Sce)* VMA intein (Asn454→Ala) and demonstrates the utility of the approach for site-specific modification of recombinant proteins [14]. EPL was used to test the hypothesis that addition of a phosphotyrosine tail to the C terminus of the protein tyrosine kinase Csk (C-terminal Src kinase) would, as in the Src family members [36,37], result in an intramolecular association between the phosphotyrosine and the Src homology 2 (SH2) domain within the protein. Because Csk is not naturally tyrosine phosphorylated at this position and contains 20 tyrosine residues within its sequence, the desired 450 amino acid phosphoprotein could not be generated using standard biotechnology techniques. Thus, EPL was used to ligate a short phosphotyrosyl-peptide to the C terminus of Csk, the reaction proceeding in essentially quantitative yield. Biochemical analysis of the semi-synthetic tail-phosphorylated Csk construct indicated the presence of an intramolecular phoshotyrosine–SH2 interaction. Interestingly, the Csk containing the phosphorylated tail appears to have increased kinase activity towards the physiologically relevant substrate, Lck, compared with the nonphosphorylated version of the protein [14].

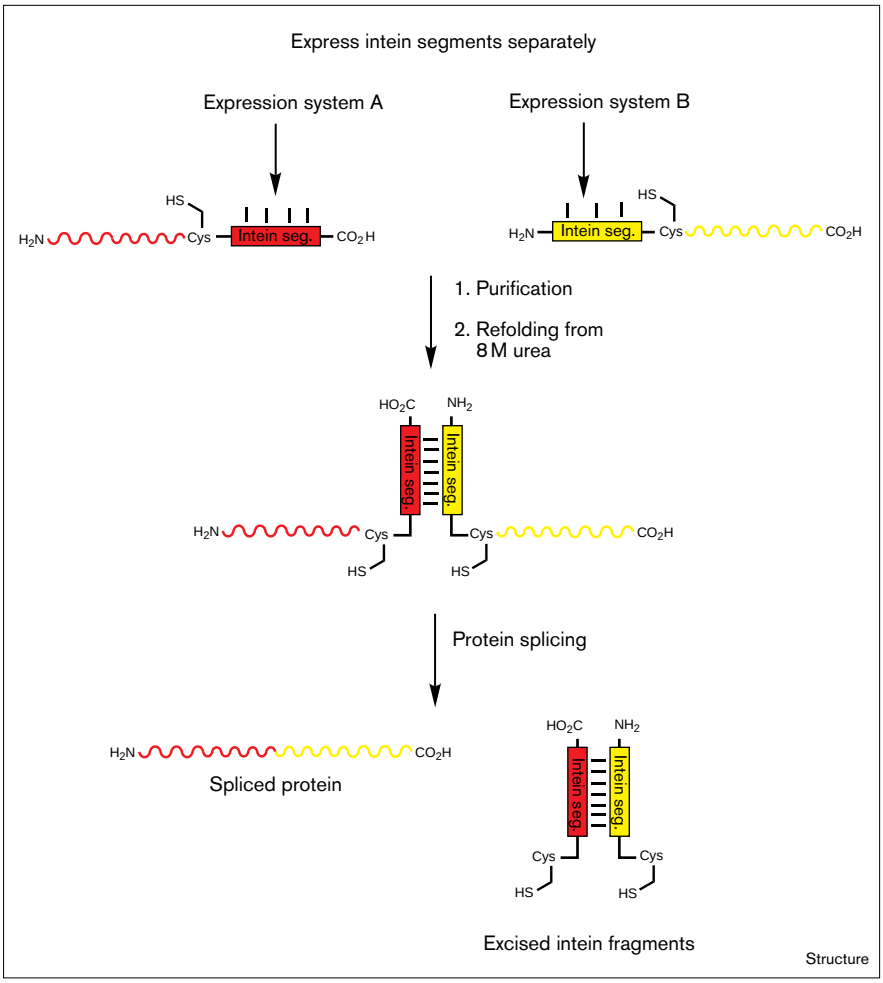
A second application of EPL, also using the commercially available *Sce* VMA intein, illustrates how the technique could prove extremely useful in studying complex biochemical machines. In this study, a fully functional semi-synthetic version of *Escherichia coli* RNA polymerase was prepared [15]. EPL was used to ligate a synthetic peptide corresponding to residues 568–600 of the $\sigma^{70}$ specificity factor to a recombinant fragment corresponding to the remainder of the $\sigma^{70}$ protein (residues 1–566). The resulting ~600 amino acid construct was shown to be fully active in initiating transcription when added to the other subunits of *E. coli* RNA polymerase, and was used to map the binding region of the bacteriophage T4 antisigma protein (AsiA) to the C-terminal region of the $\sigma^{70}$ subunit. The ability to introduce biochemical or biophysical probes into large multiprotein complexes, such as RNA polymerase, in a site-specific manner may provide new ways to study the structure and function of these complex biochemical machines. Along these lines, recent studies from our laboratory indicate that EPL can be used to introduce fluorescent probes into the preassembled ~500 kDa *E. coli* RNA polymerase complex in a site- and subunit-specific manner (TWM *et al.*, unpublished data).

In a third application, EPL has been used to prepare fully active semi-synthetic versions of bovine pancreatic ribonuclease (RNase A) and the restriction endonuclease, *Hpa* I, from *Haemophilus parainfluenzae* [16]. Both of these proteins are potentially cytotoxic to bacteria when expressed in their full-length forms, making their overexpression inefficient. The use of EPL, however, allowed inactive N-terminal fragments of the proteins to be safely expressed and then ligated *in vitro* to short synthetic peptides corresponding to the remainder of the proteins. Following refolding, the full-length semi-synthetic proteins were shown to have enzymatic properties indistinguishable from their natural counterparts. Interestingly, these semi-synthetic proteins were prepared using a slightly modified version of EPL as compared to the studies described above: a genetically engineered *Mycobacterium xenopi (Mxe)* GyrA intein was used rather than an *Sce* VMA intein, and 2-mercaptoethanesulfonic acid was used rather than thiophenol. Use of this intein/thiol combination was found to give higher yields of semi-synthetic product for these two examples. This suggests, as the authors point out, that for some targets more than one intein system may have to be tried to obtain optimal ligation results.

### Protein splicing in *trans*

The recent observation that protein splicing can be triggered *in vitro* by simply reconstituting inactive N- and

**Figure 2**




Protein splicing in *trans*. This *in vitro* strategy can be used to chemically ligate two recombinant polypeptide chains which have been expressed and purified separately. The two recombinant proteins to be linked together, represented in red and yellow, are each fused to a nonfunctional segment of an intein. When reconstituted *in vitro*, the two intein segments noncovalently associate to give a folded intein capable of supporting protein splicing (as outlined in Figure 1b).

C-terminal fragments of an intein [38–41], adds an extra dimension to the use of splicing in protein engineering. So-called *trans*-splicing is made possible because of the modular composition of the intein element; of the ~40 inteins identified to date all but two (the *Porphra purpurea* DNaB intein [42] and the *Mxe* GyrA intein [43]) contain a homing endonuclease protein module sandwiched between the functionally important flanking regions of the intein. Indeed, mutagenesis and sequence alignment studies have shown that this entire endonuclease domain can be deleted from the intein without destroying splicing activity in the resulting minimal inteins [40,43]. This finding suggested that it might be possible to obtain a folded, functional intein from two noncovalently associated intein fragments. This has in fact been shown to be the case in two systems: the Psp Pol-1 intein [38] and the *Mycobacterium tuberculosis* RecA intein [39,40]. In the process of *trans*-splicing, the two proteins to be joined are expressed as constructs with one of a pair of complementary intein fragments. Bringing these two proteins together *in vitro* under suitable conditions results in splicing (see Figure 2).

As illustrated in Figure 2, the real attraction of *trans*-splicing from a protein engineering point of view, is that it allows recombinant proteins from two different sources to be ligated together *in vitro*. As with EPL, this feature may prove to be extremely useful for obtaining cytotoxic proteins impossible to generate using standard approaches. An additional use of *trans*-splicing was recently reported by Yamazaki and coworkers who used fragments of the *Pyrococcus furiosus* PI-*PfuI* intein to link together two domains of the α subunit of *E. coli* RNA polymerase [41]. This system allowed the group to express one half of the α subunit in $^{15}N$ labeled media and the other in standard media. Consequently, the product obtained after ligation was partially labeled with $^{15}N$ allowing this domain to be selectively detected using heteronuclear two-dimensional nuclear magnetic resonance (2D NMR) experiments. As demonstrated by this work, *trans*-splicing may have enormous

implications for the study of large multidomain proteins by high-field NMR spectroscopy.

### Future outlook

One of the most important features of EPL and (potentially) *trans*-splicing is that they are technically easy to perform, a basic knowledge in recombinant protein expression is really all that is required. Moreover, in the case of EPL the necessary expression vector is commercially available and customized peptides containing N-terminal cysteine residues can be obtained from a variety of sources. How then can these approaches be used to complement the type of information already available through standard biotechnology techniques? In the case of EPL, the ability to incorporate biochemical probes (e.g. cross-linkers), biophysical probes (e.g. fluorophores and spin labels) and post-translational modifications (e.g. phosphorylated and glycosylated residues) into large proteins and protein complexes in a site-specific manner will doubtless be of enormous value in the study of these systems. Equally, the ability to obtain cytotoxic proteins either by EPL or *trans*-splicing will allow such proteins to be generated in sufficient quantities for rigorous structural and functional studies. Segmental isotope labeling, which has also been successfully performed by EPL of labeled and unlabeled recombinant proteins (TWM *et al.*, unpublished data), represents an exciting application of these technologies in the field of structural biology.

As pointed out earlier in this article, methods already exist for the generation of recombinant proteins containing an N-terminal cysteine residue. With the advent of EPL, native chemical ligation reactions can now be performed at either end of a recombinant protein. This of course allows two recombinant proteins to be directly ligated together (similar to *trans*-splicing), but it also opens the way for the insertion of synthetic peptides into recombinant proteins. The most obvious way to achieve this is to adapt the sequential native chemical ligation strategies already in use for the directed assembly of three or more synthetic peptides [27,44]. The implication of being able to specifically insert small synthetic peptide cassettes into large proteins is that the entire primary structure is made amenable to modification, and not merely the extreme N- or C-terminal regions. It is also worth noting that introduction of both an N-terminal cysteine and a C-terminal thioester within the same polypeptide chain allows intramolecular native chemical ligation reactions to be performed [45,46], a process which has been used to prepare a synthetic circular protein domain [47]. It is easy to conceive how head-to-tail cyclized recombinant peptides and proteins could be obtained in an analogous manner using EPL.

This review began with the statement that protein engineering is a multidisciplinary endeavor, requiring the interactions of many scientific fields. The novel technique of EPL continues in that tradition, combining chemistry and biology to pave new roads towards chemically engineered macromolecules.

### Acknowledgements

We thank Ming-Qun Xu for providing a copy of his manuscript prior to publication and Graham C Cotton and Francine Perler for useful discussions. This work was supported by grants from the National Institutes of Health Grant (GM55843-01; TWM), the Pew Foundation (TWM) and the National Leukemia Research Association (TWM).

### References

1. Woods, A.C., Guillemette, J.G., Pasrrish, J.C., Smith, M. & Wallace, C.J.A. (1996). Synergy in protein engineering: mutagenic manipulation of protein structure to simplify semisynthesis. *J. Biol. Chem.* **271**, 32008-32015.
2. Cornish, V.W., Mendel, D. & Schultz, P.G. (1995). Probing protein structure and function with an expanded genetic code. *Angew. Chem. Int. Ed. Engl.* **34**, 621-633.
3. Gaertner, H.F., Rose, K., Cotton, R., Timms, D., Camble R. & Offord, R.E. (1992). Construction of protein analogues by site-specific condensation of unprotected peptides. *Bioconjugate Chem.* **3**, 262-268.
4. Wallace, C.J.A. & Clark-Lewis, I. (1992). Functional role of heme ligation in cytochrome *c. J. Biol. Chem.* **267**, 3852-3861.
5. Schnölzer, M. & Kent, S.B.H. (1992). Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* **256**, 221-225.
6. Kemp, D.S. & Carey, R.I. (1993). Synthesis of a 39-peptide and a 25-peptide by thiol-capture ligations: observation of a 40-fold rate acceleration of the intramolecular O,N-acyl transfer reaction between peptide fragments bearing only cysteine protecting groups. *J. Org. Chem.* **58**, 2216-2222.
7. Rose, K. (1994). Facile synthesis of homogeneous artificial proteins. *J. Am. Chem. Soc.* **116**, 30-33.
8. Liu, C.-F. & Tam, J.P. (1994). Chemical ligation approach to form a peptide bond between unprotected peptide segments. Concept and model study. *J. Am. Chem. Soc.* **116**, 4149-4153.
9. Nefzi, A., Sun, X. & Mutter, M. (1995). Chemoselective ligation of multifunctional peptides to topological templates via thioether formation for TASP synthesis. *Tetrahedron Lett.* **36**, 229-230.
10. Gaertner, H.F., Offord, R.E., Cotton, R., Timms, D., Camble, R. & Rose, K. (1994). Chemo-enzymatic backbone engineering of proteins. *J. Biol. Chem.* **269**, 7224-7230.
11. Jackson, D.Y., Burnier, J., Quan, C., Stanley, M., Tom, J. & Wells, J.A. (1994). A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* **266**, 243-247.
12. Chang, T.K., Jackson, D.Y., Burnier, J.P. & Wells, J.A. (1994). Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl Acad. Sci. USA* **91**, 12544-12548.
13. Witte, K., Sears, P., Martin, R. & Wong, C.H. (1997). Enzymatic glycoprotein synthesis: preparation of ribonuclease glycoforms via enzymatic glycopeptide condensation and glycosylation. *J. Am. Chem. Soc.* **119**, 2114-2118.
14. Muir, T.W., Sondhi, D. & Cole, P.A. (1998). Expressed protein ligation: a general method for protein engineering. *Proc. Natl Acad. Sci. USA* **95**, 6705-6710.
15. Severinov, K. & Muir, T.W. (1998). Expressed protein ligation, a novel method for studying protein–protein interactions in transcription. *J. Biol. Chem.* **273**, 16205-16209.
16. Evans, T.C.Jr, Benner, J. & Xu, M.-Q. (1998). Semi-synthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* in press.
17. Perler, F.B., Xu, M.-Q. & Paulus, H. (1997). Protein splicing and autoproteolysis mechanisms. *Curr. Opin. Chem. Biol.* **1**, 292-299.
18. Kent, S.B.H. (1988). Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **57**, 957-989.
19. Muir, T.W. (1995). A chemical approach to the construction of multimeric protein assemblies. *Structure* **3**, 649-652.
20. Alewood, P., Alewood, D., Miranda, L., Love, S., Meutermans, W. & Wilson, D. (1997). Rapid *in situ* neutralization protocols for Boc and Fmoc solid-phase chemistries. *Methods Enzymol.* **289**, 14-29.
21. Lloyd-Williams, P., Albericio, F. & E. Giralt, (1993). Convergent solid-phase peptide synthesis. *Tetrahedron* **48**, 11065-11133.

22. Dawson, P.E., Muir, T.W., Clark-Lewis, I. & Kent, S.B.H. (1994). Synthesis of proteins by native chemical ligation. *Science* **266**, 776-779.
23. Tam, J.P., Lu, Y.-A., Liu, C.-F. & Shao, J. (1995). Peptide synthesis using unprotected peptides through orthogonal coupling methods. *Proc. Natl Acad. Sci. USA* **92**, 12485-12489.
24. Hackeng, T.M., Mounier, C.M., Bon, C., Dawson, P.E., Griffin J.H. & Kent, S.B.H. (1997). Total chemical synthesis of enzymatically active human type II secretory phopholipase A(2). *Proc. Natl Acad. Sci. USA* **94**, 7845-7850.
25. Dawson, P.E., Churchill, M.J., Ghadiri, M.R. & Kent, S.B.H. (1997). Modulation in native chemical ligation through the use of thiol additives. *J. Am. Chem. Soc.* **119**, 4325-4329.
26. Lu, W., Qasim, M.A. & Kent, S.B.H. (1996). Comparative total synthesis of turkey ovomucoid third domain by both stepwise solid phase synthesis and native chemical ligation. *J. Am. Chem. Soc.* **118**, 8518-8523.
27. Camarero, J.A., Cotton, G.C., Adeva, A. & Muir, T.W. (1998). Chemical ligation of unprotected peptides directly from a solid support. *J. Peptide Res.* **51**, 303-316.
28. Erlandson, D.A., Chytil, M. & Verdine, G.L. (1996). The leucine zipper domain controls the orientation of AP-1 in the NFAT·AP-1·DNA complex. *Chem. Biol.* **3**, 981-991.
29. Kane, P.M., Yamashiro, C.T., Wolczyk, D.F., Neff, N., Goebl, M. & Stevens, T.H. (1990). Protein splicing converts the yeast *TFP1* gene product to the 69-kD subunit of the vacuolar $H^+$-adenosine triphosphatase. *Science* **250**, 651-657.
30. Perler, F.B., *et al.*, & Belford, M. (1994). Protein splicing elements: inteins and exteins – a definition of terms and recommended nomenclature. *Nucleic Acids Res.* **22**, 1125-1127.
31. Xu, M.-Q. & Perler, F.B. (1996). The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15**, 5146-5153.
32. Shao, Y., Xu, M.-Q. & Paulus, H. (1996). Protein splicing: evidence of an N–O acyl rearrangement as the initial step in the splicing process. *Biochemistry* **35**, 3810-3815.
33. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F.B. & Xu, M.-Q. (1996). Protein splicing involving the *Saccharomyces cerevisiae* VMA intein: the steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an *in vitro* splicing system. *J. Biol. Chem.* **271**, 22159-22168.
34. Klabunde, T., Sharma, S., Telenti, A., Jacobs , W.R., Jr. & Sacchettini, J.C. (1998). Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. *Nat. Struct. Biol.* **5**, 31-36.
35. Chong, S., *et al.*, & Xu, M.-Q. (1997). Single column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271-281.
36. Sicheri, F., Moarefi, I. & Kuriyan, J. (1997). Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602-609.
37. Xu, W., Harrison, S.C. & Eck, M.J. (1997). Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595-602.
38. Southworth, M.W., Adam, E., Panne, D., Byer, R., Kautz, R. & Perler, F.B. (1998). Control of protein splicing by intein fragment reassembly. *EMBO J.* **17**, 918-926.
39. Shingledecker, K., Jiang, S.-Q. & Paulus, H. (1998). Molecular dissection of the *Mycobacterium tuberculosis* RecA intein: design of a minimal intein and of a *trans*-splicing system involving two intein fragments. *Gene* **207**, 187-195.
40. Mills, K.V., Lew, B.M., Jiang, S.-Q. & Paulus, H. (1998). Protein slicing in *trans* by purified N- and C-terminal fragments of the *Mycobacterium tuberculosis* RecA intein. *Proc. Natl Acad. Sci. USA* **95**, 3343-3548.
41. Yamazaki, T., *et al.*, & Nakamura, H. (1998). Segmental isotope labeling for protein NMR using peptide splicing. *J. Am. Chem. Soc.* **120**, 5591-5592.
42. Pietrokovski, S., (1996). A new intein in cyanobacteria and its influence for the spread of inteins. *Trends Genet.* **12**, 287-288.
43. Telenti, A., Southworth, M., Alcaide, F., Daugelat, S., Jacobs, W.R.Jr. & Perler, F.B. (1997). The *Mycobacterium xenop*i GyrA protein splicing element: characterization of a minimal intein. *J. Bacteriol.* **179**, 6378-6382.
44. Muir, T.W., Dawson, P.E. & Kent, S.B.H. (1997). Protein synthesis by chemical ligation of unprotected peptides in aqueous solution. *Methods Enzymol.* **289**, 266-298.
45. Zhang, L. & Tam, J.P. (1997). Synthesis and application of unprotected cyclic peptides as building blocks for peptide dendrimers. *J. Am. Chem. Soc.* **119**, 2363-2370.
46. Camarero, J.A. & Muir, T.W. (1997). Chemoselective backbone cyclization of unprotected peptides. *J. Chem. Soc. Chem. Commun.* 1369-1370.
47. Camarero, J.A., Pavel, J. & Muir, T.W. (1998). Chemical synthesis of a circular protein domain: evidence for folding-assisted cyclization. *Angew. Chem. Int. Ed.* **37**, 347-349.